

Last year's accomplishments:

1. Multiple credential support:

GridFTP users may not always use the same credential between any 2 GridFTP server endpoints, but globus-url-copy CLI does not support the ability to associate a credential to be used for a specific endpoint. This feature adds endpoint specific control channel authentication/authorization.

2. GridFTP on cygwin:

Advanced Photon Source (APS) at Argonne uses GridFTP for automated data movement between the beamline data acquisition machine to its HPC cluster. The data acquisition machine is a windows machine. CDIGS team provided a distribution of GridFTP for cygwin to APS. Using this version of GridFTP, APS is getting data rates of about 110 MB/s, a 5x improvement over the data rates they have been getting earlier using a windows native protocol.

3. Transfer resume capability in globus-url-copy:

Globus-url-copy (commonly used GridFTP client) provides ways to recover from server and connection failures as long as globus-url-copy is not dead. globus-url-copy can not recover from its own failures. A number of users (including many teragrid and OSG users) use globus-url-copy to transfer data sets involving lot of files with complex directory structures. When globus-url-copy fails in the middle of a transfer, it is tedious for them to manually figure out the missing files and transfer those. The CDIGS team has developed a transfer resume capability for globus-url-copy by storing minimal information in client's local filesystem.

4. Developed alternate proposal for firewall traversal – single port gridftp:

The initial design the CDIGS team proposed in the previous years for addressing the firewall issues with GridFTP requires significant modifications to the protocol. This year CDIGS team has developed an alternate proposal to address the firewall problem. The proposal is to develop a single port GridFTP – use a single static port for both control and data channel. The proposal is available at <http://jira.globus.org/browse/GRIDFTP-15>

5. Documentation improvements:

Created a quick start guide for SSH GridFTP

Created FAQs for GridFTP

Created documentation on how to limit bandwidth for outgoing traffic on a given port range using networking features of the Linux

The CDIGS team has developed an initial draft of best practices guide on

- Protecting your GridFTP server from unexpectedly heavy usage
- Increasing your GridFTP server's ability to satisfy heavy usage
- Providing a high-availability GridFTP service

6. Fixed a security issue in GridFTP in a timely manner

Gridftp was intermittently giving users access to groups that root is in but the user is not. As soon as this issue is uncovered, the CDIGS team diligently worked on it and provided a fix in a timely manner.

7. Provided extensive help to the following user communities:

- Structural Biology on the Grid from Harvard and the ATLAS community to fix a firewall related problem
- LIGO on a number of issues (eg. Issue of zombie GridFTP processes hanging around after the transfer is done, Client deadlock problem on solaris 10)
- ESG on number of issues (eg. Creating an authz callout for GridFTP, fix a firewall related problem)
- VPAC from Australia on using pipelining to improve the performance lots of small files transfers, using ssh GridFTP and also using UDT
- Indian Space Research Organization using GridFTP for "Satellite-based Grid Technology Development" where the transmission delay is the order of 500-1000 ms and Bit Error rate is of order 10^6 . Based upon different satellite conditions, they had to set GridFTP Parallel streams up to 800.
- Resolved a ssh GridFTP problems with specific versions of OpenSSH for NASA

Quote from a user:

"With great effort from CDIGS team and LLNL network people, GridFTP connections to outside from pcmdi3.llnl.gov is now resolved." – Alex Sim from LBNL

8. Press releases and news articles:

"Argonne technology enables high-speed data transfer" Argonne press release, June 2009 (http://www.anl.gov/Media_Center/News/2009/news090617.html)

"GridFTP accelerates traffic on the World Wide Grid" International Science Grid This Week, July 2009 (<http://www.isgtw.org/?pid=1001944>)

9. Relevant publications:

"GridFTP Multilinking", John Bresnahan, Michael Link, Rajkumar Kettimuthu and Ian Foster, Proceedings of the 2009 TeraGrid Conference, Arlington, VA, June 2009

"UDT as an Alternative Transport Protocol for GridFTP", John Bresnahan, Michael Link, Rajkumar Kettimuthu and Ian Foster, Proceedings of the 7th International Workshop on Protocols for Future, Large-Scale and Diverse Network Transports (PFLDNeT 2009), Tokyo, Japan, May 2009

10. Tutorials:

"GridFTP for Administrators," Argonne National Laboratory, Argonne, IL, September 2009.

"GridFTP for Users," Argonne National Laboratory, Argonne, IL, September 2009.

11. Talks:

"GridFTP: Challenges in Bulk Data Movement," Raj Kettimuthu, Institute for Scientific Computing, University of Vienna, Vienna, Austria, Sep 2009

"High Performance Data Movement using GridFTP," Raj Kettimuthu, Center for Development of Advanced Computing, Bangalore, India, May 2009

Next year goals:

1. Data channel security for SSH GridFTP
2. Automate GridFTP build for cygwin
3. Single/dual port GridFTP
4. Develop methods to do data channel authentication when the users use different credentials to authenticate with the source and destination GridFTP servers